DYNAMIC RATE CONTROL FOR JPEG 2000 TRANSCODING

Derek Schwenke[†], Anthony Vetro[†], Toshihiko Hata[‡], Naoki Kuwahara[‡]

[†] Mitsubishi Electric Research Labs, Cambridge, MA 02139, USA
[‡] Mitsubishi Electric Corporation, Amagasaki, Hyogo 661-8661, Japan

ABSTRACT

This paper describes a rate control algorithm for a streaming video system that dynamically transcodes stored JPEG 2000 frames. The proposed algorithm is designed to improve overall quality over a static rate control method by increasing bandwidth utilization, while satisfying buffer constraints and maintaining consistent quality over time. Simulation results confirm the effectiveness of the proposed algorithm in terms of both objective measures and subjective evaluation.

1. INTRODUCTION

The JPEG 2000 standard [1] is becoming an increasing popular coding format for applications that require scalable transmission of images or video. In our prior work, we have developed an object-aware video surveillance system based on JPEG 2000 that is not only smart and friendly for users, but allows for transmission of the scene over limited bandwidth networks [2]. In this system, an image sequence is encoded and stored as a JPEG 2000 bitstream, and then the stored images are efficiently transcoded in the compressed-domain using a low-complexity adaptation technique that replaces data packets corresponding to higher quality layers with empty packets. In one particular streaming mode, the ROI (Regions-of-Interest) are transcoded with higher quality than the background to satisfy network constraints.

This paper considers the problem of rate allocation to each frame. One straightforward method, which will be referred to as static rate control and is used as a reference in this work, is to allocate an equal amount of rate to each frame based on available channel bandwidth. The obvious drawback of this method is that it is not adaptive to the scene contents. Also, since there is a fixed set of rate points that could be achieved by the transcoder, which depends on the rate allocated to each quality layer and other transcoding parameters such as output resolution level and ROI, it is very likely that the available bandwidth is not fully utilized.

We propose a dynamic rate control technique that is adaptive to scene contents. The proposed rate allocation is designed to improve overall quality over the static rate control method by utilizing more of the available bandwidth, while satisfying buffer constraints and maintaining consistent quality over time.

The rest of this paper is organized as follows. In the next section, we provide a brief overview of our system. In section 3, the dynamic rate control algorithm is presented. Experimental results are described in section 4, including both objective and subjective evaluation. Concluding remarks are given in section 5.

2. SYSTEM OVERVIEW

For completeness, we provide an overview of the key components of our transcoding system [2]. As shown in Fig 1, input image frames are encoded by a standard JPEG 2000 encoder to yield the input code stream. The input image sequence is also subject to an object tracking algorithm to yield the ROI information, which is specified by a rectangular bounding box surrounding the extracted object. In our system, the code stream and resulting ROI information are stored. To satisfy network and display constraints, transcoding is performed to yield the output code stream. In the following, we describe three main components of our JPEG 2000 transcoder including data analysis, ROI transcoding and quality control.

The data analysis module is responsible for extracting indexing information about the structure of the code stream. It is essentially a low-complexity parser that analyzes the packet header for each quality layer, resolution level and component. A multiple-dimensional array is used to store the packet information, which indicates the byte position, header length and body length for each packet. Since this partial decoding operates on the packet header only without performing entropy arithmetic decoding for code blocks, the computational complexity is very low.

Our transcoder supports reduction of spatial resolution and quality layers. We focus mainly on quality layer reduction considering the ROI information. Given a set of ROI coordinates, we perform ROI transcoding by replacing packets at high quality layers that are associated with the background of the scene with empty packets as defined by the JPEG 2000 standard. This is an effective method for reducing the rate of the overall code stream while retaining the quality of important objects and keeping the complexity low.
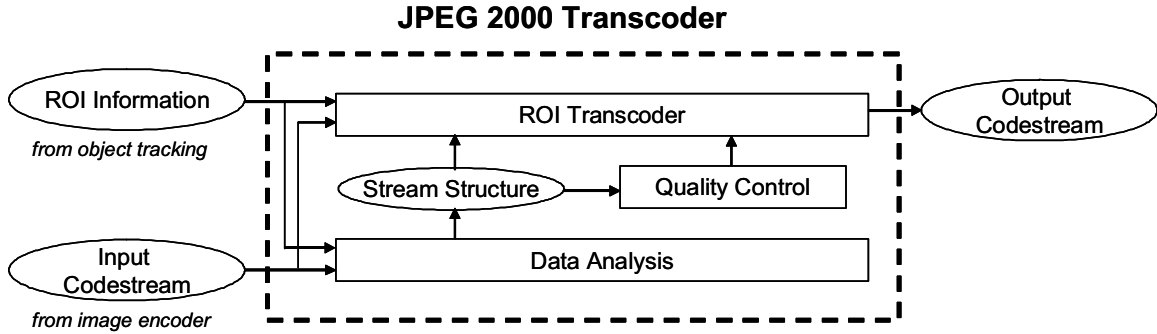
Figure 1. Overview of JPEG 2000 transcoding system.

The number of quality layers for the background and ROI are determined by the quality control module. In our previous work [2], the quality layers were set manually. In the next section, we describe a dynamic rate control algorithm that determines the quality layers based on target rate, buffer occupancy and ROI information.

3. DYNAMIC RATE CONTROL

The proposed dynamic rate control algorithm determines the rate allocation for the current frame based on the target rate, buffer occupancy and ROI information. Given the bytes allocated to a frame, the transcoder determines quality layers for background and ROI. In the following, we describe the variable rate allocation, a frame skipping technique, as well as a quality stabilization algorithm.

3.1 Variable Rate Allocation

In the static rate control method, a fixed rate, $T_f = R/F$, is allocated to each frame, where R is the target rate and F is the output frame rate. To avoid overshooting the target rate, the quality layers in the transcoded output are chosen so as not to exceed the given budget. In our current system, we choose the quality layers for background and ROI in a systematic manner based on byte counts from the data analysis. We first assign the minimum quality to the background and ROI. The ROI quality is then successively increased. Finally, additional quality layers are added to the background. The main drawback of this static rate allocation approach is that it will typically underutilize the available bandwidth for a given stream because the quality layers can only provide a discrete set of rate points.

In the proposed rate control algorithm, we allocate rate non-uniformly to each frame and introduce a buffer to absorb the variations in allocated rate to each frame. The rate allocation to each frame is determined according to the following:

$$T_v = T_{max} \cdot \max\left[0, \min(1, 1 - \alpha^2)\right],\qquad(1)$$

where T_{max} sets the upper limit on the variable rate assigned to any frame and is $2T_f$ in our current system, and α is a buffer occupancy parameter that is a function of the buffer occupancy, B, the buffer size, B_s, and a safety margin, γ, with typical values in the range [0.05, 0.25]. The buffer occupancy parameter is given by,

$$\alpha = \frac{B}{B_s \cdot (1 - \gamma)}.\qquad(2)$$

When the buffer occupancy is near the upper margin, α tends towards unity, and a lower rate will be allocated to the current frame. Higher rate is allocated to the current frame when the buffer occupancy is near empty. In the next subsection, we will see how this behavior plays an important role in balancing the spatio-temporal quality trade-off when frame skipping is employed.

3.2 Frame Skipping

When frame skipping is enabled, periodic frames with no ROI defined may be skipped. The rationale behind this strategy is twofold. First, we aim to empty the buffer when there is no ROI to allow greater bandwidth for future frames that contain ROI. Second, we aim to improve the quality of the non-ROI image, which is possible since we could assign more bytes to an image sequence with a reduced frame rate.

With this strategy, frames are skipped to drive the buffer level towards its lower margin. When the buffer reaches this level, frames will no longer be skipped, and since the buffer is nearly empty, these frames are allocated a rate close to T_{max}.

Table 1. Experimental results comparing SRC with various DRC configurations.

Configuration	Skip	BWU	ΔP0	ΔP1	Avg0	Avg1	MSE
SRC	N/A	0.70	86	99	2.0	4.1	61.2
DRC	N/A	1.00	606	13	3.2	4.9	40.1
DRC + W(4)	N/A	1.00	155	7	3.1	4.9	40.6
DRC + Skip(7)	493	0.84	34	1	3.7	5.0	26.1
DRC + Skip(7) + W(4)	493	0.84	25	1	3.8	5.0	26.0

To state the skip condition more precisely, a frame is skipped when the following condition is true:

$$(\alpha > \gamma) \quad \& \quad (\tau < \tau_{\max}) \tag{3}$$

where τ is the interval of successive non-ROI frame skips, and $\tau_{\max}$ is the maximum frame skip interval.

3.3 Quality Stabilization

The key objective of the quality stabilization is to establish a period in which the quality layers will be held stable, thereby avoiding unnecessary oscillation or frequent changes in quality. Depending on the available buffer size, the typical window period, ω, will be several frames.

Let Q_p denote the set of determined quality layers for the previous frame, Q_i the set of quality layers for the current frame i with rate allocated according to eqn. (1), and ω_c be a window counter that is reset when either the counter reaches the window period or a new set of quality layers for the current frame are determined. With quality stabilization enabled, the set of quality layers would be assigned according to:

$$Q = \begin{cases} Q_p & B_s \cdot \gamma < B < B_s \cdot (1-\gamma) \& (\omega_c < \omega) \\ Q_i & otherwise \end{cases}. \tag{4}$$

With the above, the previous set of quality layers will be used for the current frame when the buffer is not in danger or overflow or underflow and the window counter is less than the window period.

4. EXPERIMENTAL RESULTS

To confirm the effectiveness of the proposed dynamic rate control algorithm, we perform a number of experiments with varying configurations and buffer size. We evaluate both objective and subjective quality and use the static rate control method as a benchmark.

For the purpose of this study, we define the following objective measures:

- BWU: bandwidth utilization defined as the ratio of transcoded output bits to the target rate
- ΔP0: number of changes in background quality
- ΔP1: number of changes in ROI quality
- Avg0: average background quality
- Avg1: average ROI quality

As one would expect, achieving higher bandwidth utilization will generally increase overall quality. Also, minimizing the fluctuation in quality layers over time also has a positive impact on perceptual quality. It is noted that the average ROI quality is not as relevant as the average background quality since the ROI typically receives high quality regardless of the rate control method or algorithms used. In addition to the above metrics, we also report the number of frames skipped and MSE. It is noted that the MSE for skipped frames is computed assuming a zero-order hold, i.e., based on the previously coded frame.

As input, we use an image sequence with 1467 frames and at a frame rate of 7.5fps. Each frame of the image sequence is a full color image (4:4:4) that is JPEG 2000 encoded with 5 quality layers and 6 resolution levels with LRCP progression. The precinct sizes are set as 64x64, 32x32, 16x16, 8x8, 4x4 and 4x4 and rate for each quality layer is 0.125, 0.0625, 0.0625, 0.25 and 0.5. The overall bit-rate of this input code stream is 2.3Mbps.

In our first experiment, we test the effectiveness of the proposed rate control components. The input code stream is transcoded to a target bit-rate of 800 kbps using the following transcoding methods: static rate control (SRC), Dynamic Rate Control (DRC) with variable rate allocation, DRC with ω=4 for quality stabilization, DRC with $\tau_{\max}$=7 for frame skipping, and DRC with both quality stabilization and frame skipping enabled. In all simulations for DRC in this experiment, the buffer size is set to 1MB.

Table 2. Experimental results comparing DRC + W(4) + Skip(7) with various buffer sizes.

Buffer	Skip	BWU	ΔP0	ΔP1	Avg0	Avg1	MSE
1MB	493	0.84	25	1	3.8	5.0	26.1
512KB	475	0.76	46	3	3.5	4.9	30.1
256KB	451	0.69	71	15	3.3	4.9	34.6
128KB	430	0.64	136	93	3.0	4.1	37.6
64KB	419	0.61	172	150	2.9	3.4	38.6

The results of this first experiment are summarized in Table 1. From the table, we observe that the bandwidth utilization and quality of both background and ROI using SRC is relatively low compared with DRC. While the overall quality of DRC is clearly higher than that of the SRC method in terms of quality layers and MSE, the quality of the background using DRC fluctuates significantly. Such oscillations in quality have a notable impact on quality for certain segments of the video. With the proposed quality stabilization algorithm, these fluctuations can be controlled with minimal change to the overall average quality. Finally, with the frame skipping enabled, we see another moderate increase in quality and fewer fluctuations in quality.

From the data, we find that frame skipping accounts for the majority of gains observed for this particular sequence. This is likely due to the relatively high percentage of non-ROI frames in the test sequence, which is a typical surveillance video. Larger differences between the skip only and skip with quality stabilization could be expected for sequences with a higher percentage of ROI frames.

In our second set of experiments, we investigate the impact of buffer size on the efficiency of the dynamic rate control algorithm. Generally speaking, larger buffers not only require more memory in a device but also increase delay. Depending on the application, limited buffers or strict requirements on the delay may be imposed. Using the same image data and target bit-rates, the DRC algorithm with quality stabilization and frame skipping is simulated with varying buffer sizes from 1MB to 64KB.

The results of the second experiment are summarized in Table 2. As expected, we see a slight decline in performance with reduced buffer sizes. With smaller buffer sizes, we observe that the bandwidth utilization is decreased and hence the average quality becomes lower. Reduced buffer sizes also constrain the effect of frame skipping, i.e., reducing the average number of bits for a coded frame and lowering overall quality. It is noted that even with reduced buffer sizes, the DRC method still outperforms SRC in terms of average overall quality. The

most significant gains will be obtained with larger buffer sizes though.

Extensive subjective evaluation has been carried out. The results reveal that the proposed DRC algorithm offers substantial improvement over SRC when either quality stabilization and/or frame skipping are enabled. Without at least one of these options, frequent fluctuations in quality occur contributing to an overall decrease in subjective quality. Furthermore, for this particular sequence tested, it has been found that DRC with quality stabilization and a large buffer size is subjectively similar to DRC with frame skipping and small buffer size. Therefore, the skip only option is preferred for low delay applications.

5. CONCLUDING REMARKS

This paper presented a dynamic rate control algorithm for JPEG 2000 transcoding. The algorithm allocates rate to each frame in an image sequence based on target rate, buffer occupancy and ROI information. The key components of the proposed rate control algorithm include variable rate allocation, frame skipping and quality stabilization. The benefits of these components have been studied and it has been shown that the proposed algorithm significantly outperforms the reference static rate control method.

We believe there is still opportunity to improve these results further, especially for applications that require a limited buffer size. Maximizing the perceptual quality considering quality fluctuation and frame skipping in non-background frames is another topic to explore further.

REFERENCES

[1] ISO/IEC 15444-1, "Information technology – JPEG 2000 image coding system – Part 1: Core coding system," 2000.

[2] T. Hata, N. Kuwahara, T. Nozawa, D. Schwenke and A. Vetro, "Surveillance System with Object-Aware Video Transcoder," *Proc. Workshop on Multimedia Signal Processing*, Shanghai, China, Nov 2005.